New insights in methodology of screening for cervical cancer

Wang, Rong

Chapter 4

Discovery of new methylation markers to improve screening for cervical intraepithelial neoplasia grade 2/3

Boers A.[1], Wang R.[1], van Leeuwen R.W.[1], Klip H.G.[1], de Bock G.H.[2], Hollema H.[3], van Criekinge W.[4], de Meyer T.[4], Denil S.[4], van der Zee A.G.J.[1], Schuuring E.[3], Wisman G.B.A.[1]

[1] Department of Gynecologic Oncology, University of Groningen, University Medical Center Groningen, the Netherlands

[2] Department of Epidemiology, University of Groningen, University Medical Center Groningen, the Netherlands

[3] Department of Pathology, University of Groningen, University Medical Center Groningen, the Netherlands

[4] Department of Molecular Biotechnology, Ghent University, Ghent, Belgium

Submitted for publication

Abstract

Aims: To identify new methylation markers for high-grade cervical intraepithelial neoplasia (CIN2/3) using innovative genome-wide methylation analysis and to assess their diagnostic performance in cervical scrapings.

Methods: Enrichment and capturing of methylated DNA from normal cervices and CIN2/3 lesions followed by next-generation sequencing (MethylCap-Seq) was performed to identify differential methylation regions (DMRs). The top 15 highest ranking differentially methylated genes were selected and validated by MSP in two steps: on the same DNA samples as used for MethylCap-Seq and on DNA samples from an independent patient cohort with (pre)malignant cervical neoplasia. For further diagnostic evaluation, the best differentiating methylation markers were tested with quantitative MSP (QMSP) in cervical scrapings from 2 cohorts: 1) cervical carcinoma vs. healthy controls and 2) patients referred from population-based screening with an abnormal Pap smear in whom HPV status was determined.

Results: With genome-wide MethylCap-Seq, 176 DMRs comprising 163 genes were identified. After verification and validation of the top 15 genes with MSP, 9 genes showed significant differential methylation in normal cervices versus CIN2/3 lesions ($p<0.05$). Subsequently, methylation levels of 8/9 genes were significantly higher in carcinoma compared to normal scrapings. For all 8 genes methylation levels increased with the severity of the underlying histological lesion in scrapings from patients with an abnormal Pap smear. In addition to the 8 new genes, also our previous four-gene panel (*C13ORF18, JAM3, EPB41L3* and *TERT*) was analyzed. The best combination of genes (*C13ORF18/JAM3/AL590705.4)* revealed sensitivity (74%) for CIN2+ comparable to hrHPV testing (79%), while specificity was significantly higher (76% vs 46%, $p\leq0.05$) in a triage setting after a positive Pap smear test in population-based screening.

Conclusion: We identified new CIN2/3 specific methylation markers using a genome-wide DNA methylation analysis. The diagnostic performance of our new methylation panel shows comparable sensitivity to hrHPV testing for CIN2+, but with higher specificity to prevent referral for unnecessary colposcopy. The next step before implementation in primary screening programs will be validation in population-based cohorts.

Introduction

Cervical cancer is characterized by a well-defined pre-malignant phase, cervical intraepithelial neoplasia (CIN). Identification of these CIN lesions by population-based screening programs and their subsequent treatment has led to a significant reduction of the incidence and mortality of cervical cancer [1,2]. Cytology-based testing of cervical smears is the most widely used cervical cancer screening method, but is not ideal, as the sensitivity for detection of CIN2 and higher (CIN2+) is only ~55% [3-5]. Cervical carcinogenesis is highly associated with high-risk human papillomavirus (hrHPV) [6]. Large randomized-controlled trials have shown that the sensitivity of hrHPV testing is significantly higher than cytology testing [4,7-10]. However, the specificity of hrHPV testing, especially in a young screening population is relatively low [3,11-13], which may lead to unnecessary referrals to the gynecologist, anxiety in the false-positive women, and higher costs for the health-care system. Finally, in the near future the prevalence of CIN and cervical cancer will probably decrease in countries that have introduced primary prevention with hrHPV vaccination. With this decrease in prevalence, the positive predictive value of the current screening tests will by definition decrease [14]. Therefore, other objective biomarkers with both high sensitivity as well as high specificity are needed as new screening tools for cervical cancer.

Different DNA methylation patterns in normal versus (pre)malignant lesions represent excellent targets for diagnostic approaches based on methylation specific PCR (MSP). Promoter hypermethylation of tumor suppressor genes is an early event in cervical carcinogenesis and consequently hypermethylation analysis can be especially relevant for the early detection of cervical neoplasia [15-17]. Assessment of methylation markers in cervical scrapings for the detection of CIN and cervical cancer is feasible [17-23], but finding methylation markers with both high sensitivity as well as high specificity remains a challenge. Through the years gradually more sophisticated approaches have been developed to identify new methylation markers on a genome-wide scale [24]. Amidst comparable studies from other groups we have previously reported our experience with pharmacological unmasking of the promoter region combined with re-expression as analyzed by microarrays, high-throughput quantitative methylation specific PCR (QMSP) on an OpenArray platform and methyl-

DNA immunoprecipitation followed by microarray analysis (MeDIP), resulting in the discovery and validation of the genes *C13ORF18, JAM3, EPB41L3* and *TERT* [21,22,25]. The diagnostic performance of these genes showed sensitivities for detecting CIN2+ in a hrHPV positive population between 43%-71% and specificities between 89%-100% [21]. However, our strategies for discovering new methylation markers so far were based on the difference between cancer and normal tissue resulting in markers with high sensitivity for carcinoma, but with too low sensitivity for detecting CIN2/3 lesions. In our MeDIP study DNA methylomes of normal and CIN3 lesions were analyzed [25]. However, a disadvantage of this technique is that it primarily recognizes bulk quantities of highly methylated repetitive DNA, resulting in less specificity. New and more specific innovative genome-wide methylation analysis of DNA from CIN2/3 lesions versus normal cervical tissue should result in (new) CIN2/3 sensitive and specific methylation markers. Methylated-CpG island recovery assay uses antibody-coupled methyl-binding domain (MBD2) proteins to specifically purify methylated DNA [26]. The higher affinity of the MBD2 complex for double-stranded CpG-methylated DNA results in a higher enrichment for methyl DNA sequences as compared to MeDIP analysis. Next-generation-sequencing then shows the identified novel methylated regions (MethylCap-Seq). After identification of novel methylation markers for (pre)malignant cervical neoplasia through this approach, validation and diagnostic evaluation of these newly found markers can be performed.

The aim of the present study was 1) to identify new methylation markers that can differentiate between normal cervices and CIN2/3 lesions using MethylCap-Seq and 2) to validate the diagnostic performance of the newly found methylation markers in cervical scrapings by QMSP.

Patients and Methods

General strategy

To characterize the DNA methylome of CIN2/3 lesions and to identify new CIN2 or higher (CIN2+) methylation markers, we applied the following strategy (see figure 1):

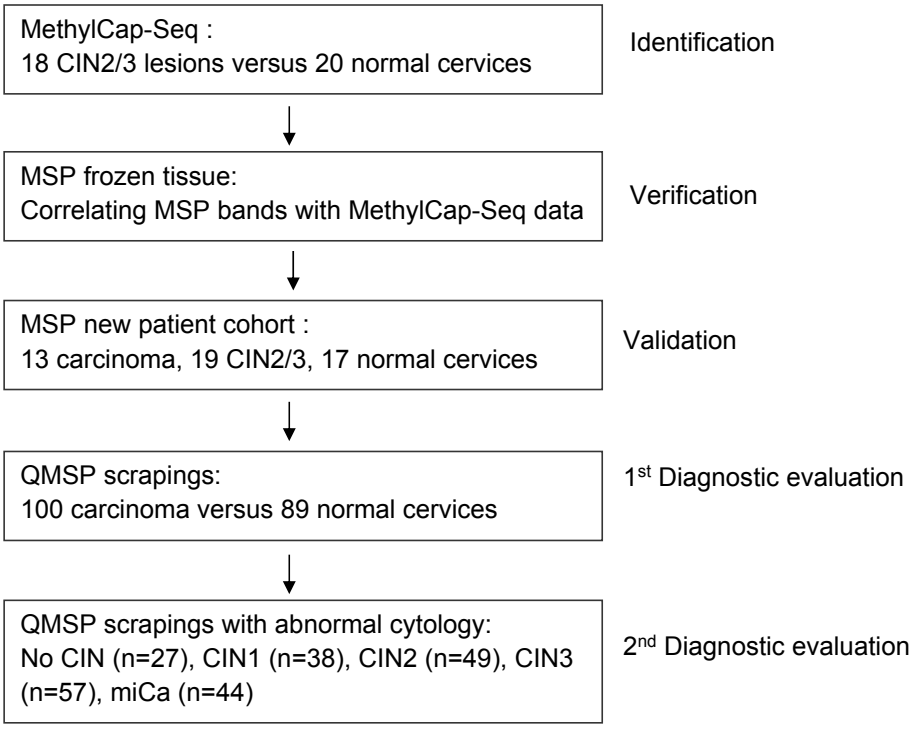

Figure 1: Flow scheme for the identification of new CIN2+ methylation markers

First, methylated DNA was enriched using MBD2 proteins with subsequent paired-end sequencing (MethylCap-Seq) on DNA isolated from fresh-frozen macro-dissected epithelial tissue of 18 CIN2/3 lesions (6 CIN2 and 12 CIN3), 20 normal cervices and two pools of leukocyte DNA of healthy volunteers. In order to identify differential methylated regions (DMRs), we retrieved the reads of promoter and exon regions. We selected methylation markers that showed significant differences between the normal and CIN2/3 cervices, while also the leukocyte count had to be low, to prevent false-positive results. Markers were ranked on high specificity (no methylation in the normal cervices) and high sensitviy (methylation in CIN2/3 lesions). For the highest ranking top15 genes, methylation specific PCR (MSP) primers were designed and methylation patterns were verified on the same DNA, which originally was used for MethylCap-Seq. This first validation step enabled verification of MethylCap-Seq data by correlating MSP band intensity with the number of reads

from the MethylCap-Seq. In the second validation step high prevalence of methylation in the CIN2/3 lesions and no methylation in the normal cervices was analyzed by MSP analysis on DNA isolated from a completely independent cohort of patients (cervical cancer (n=13), CIN2/3 lesions (n=19) and normal cervices (n=17)). DNA was isolated from macro-dissected formalin fixed paraffin embedded (FFPE) epithelial tissue.

Finally, diagnostic evaluation of the newly discovered methylation markers was performed by QMSP on cervical scrapings. First, we tested the methylation ratios of new biomarkers on a large series of randomly selected scrapings from cervical cancer patients (n=100) and a similar age group of healthy controls (n=89). Secondly, the potential of the new methylation markers as a diagnostic tool was evaluated in a large series of scrapings (n=215) of randomly selected patients, referred with an abnormal Pap smear at population-based screening. Histology was used as the reference standard.

Patient samples

All patients referred to the outpatient clinic of the University Medical Center Groningen (UMCG) with cervical cancer or an abnormal Pap smear at population-based screening are routinely asked to participate in our ongoing 'Methylation study' which has been approved by the Institutional Review Board (IRB) of the UMCG. Cervical tissue, scrapings and clinicopathologic data are prospectively collected and stored in our tissue bank. Within our Methylation study tissue samples, scrapings and clinicopathologic data from normal cervices are also collected from patients planned to undergo a hysterectomy for non-malignant reasons. All cervical tissue that was used for the normal control group was judged as histopathological normal. Patients referred with cervical cancer are staged according to the FIGO criteria with pelvic examination and biopsies under general anaesthesia. Cervical scrapings from both groups (cervical cancer staging and benign gynecologic surgery) were collected before surgery under general anaesthesia. All patients referred with an abnormal Pap smear at population-based screening underwent an additional Pap smear prior to colposcopy specifically for this study. At colposcopy, biopsies and/or Large Loop Excision of the Transformation Zone (LLETZ) were performed. The tissue samples

were scored by an experienced gynecologic pathologist and the histological classification was used as the reference standard. If no interference with routine diagnostic evaluation was anticipated, specimens from the CIN lesions were retrieved and stored at -80 °C. Clinicopathological data were retrieved from patient files and stored in our large anonymous password-protected institutional Gynecologic Oncology database. All patients gave written informed consent.

For the frozen tissue samples used in de MethylCap-Seq analysis, the median age of the CIN2/3 patients was 35 years (IQR 30-39) and for the patients with normal cervices 43 years (IQR 41-44). For the independent cohort of patients with FFPE samples, the median age of the CIN2/3 patients was 37 years (IQR 34-41), for the patients with normal cervices 43 years (IQR 40-44) and for the cervical cancer patients 49 years (range 42-54). For the cervical scrapings the median age of cervical cancer patients was 50 years (IQR 39-64) and for the patients with normal cervices 47 years (IQR 43-53). The stage of cervical cancer patients was: 1 (1%) FIGO stage IA1, 31 (31%) FIGO stage IB1, 18 (18%) FIGO stage IB2, 21 (21%) FIGO stage IIA, 17 (17%) FIGO stage IIB, 1 (1%) FIGO stage IIIA, 8 (8%) FIGO stage IIIB and 3 (3%) FIGO stage IV. Histological classification of the cervical cancer patients was: 70 (70%) squamous cell carcinoma (SCC), 21 (21%) adenocarcinoma (ADC), 3 (3%) adenosquamous (ASC) and 6 (6%) undifferentiated carcinoma. The median age of the patients referred with an abnormal Pap smear was 37 years (IQR 32-43). The histological classifications of these patients were: 27 without CIN, 38 CIN1, 49 CIN2, 57 CIN3 and 44 miCa (29 SCC, 12 ADC, 3 ASC). The Pap smears were classified according to the Papanicolaou system.

From all frozen tissue samples used for MethylCap-Seq and the FFPE samples, 10 µm tissue sections were cut and macrodissection was performed to enrich for epithelial cells. Before and after cutting a hematoxylin and eosin slide was made to check presence of epithelial cells. Cervical scrapings were collected in 5 ml ice-cold phosphate buffered saline (PBS: 6.4 mM NA_2HPO_4; 1.5 mM KH_2PO_4; 0.14 M NaCl; 2.7 mM KCl) and kept on ice until further processing. Of these 5 ml cell suspension, 1 ml was used for cytomorphological assessment. The remaining 4 ml was centrifuged and the cell pellet was suspended in 1 ml TRAP wash buffer and divided in 4

fractions. Two fractions were stored as dry pellet at -80°C for DNA isolation as described previously [21].

DNA isolation

Tissue slides from FFPE tissue were deparaffinized using 100% xylene followed by 100% ethanol [17]. Genomic DNA from fresh-frozen macro-dissected samples and cervical scrapings was isolated by standard overnight 1% SDS and Proteinase K treatment, salt-chloroform extraction and isopropanol precipitation as described previously [21]. DNA pellets were washed with 70% ethanol and dissolved in 150 µl TE^{-4} (10 mM Tris/HCL; 0.1 mM EDTA, pH 8.0). Genomic DNA was amplified in a multiplex PCR according to the BIOMED-2 protocol, to check the DNA's structural integrity [27]. For the MethylCap-Seq samples, DNA quantity was measured using Quant-iT™ PicoGreen® dsDNA Assay Kit according to manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). For cervical scrapings DNA concentrations and 260/280 ratios were measured using the Nanodrop ND-1000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA). A 260/280 ratio of >1.8 was required for all samples.

Methylated-CpG island DNA capturing followed by next-generation sequencing (MethylCap-Seq)

Methylated DNA fragments were captured with methyl-binding domains using the MethylCap kit according to manufacturers instructions (Diagenode, Liège, Belgium). The kit consists of the methyl binding domain (MBD) of human MeCP2, as a C-terminal fusion with Glutathione-S-transferase (GST) containing an N-terminal His6-tag. Before capturing, DNA samples (500 ng) were sheared to a size range of 300-1000 bps using a Bioruptor™ UCD-200 (Diagenode, Liège, Belgium) and fragments of ~300 bp were isolated. Leukocyte DNA of 4 healthy controls were included in 2 sets of 2 samples. Captured DNA was paired-end-sequenced on the Illumina Genome Analyzer II platform according to protocol (Illumina, San Diego, CA, USA). Results were mapped on the nucleotide sequence using Bowtie software [28], visualized using BioBix' H2G2 browser (http://h2g2.ugent.be/) and processed using the human reference genome (NCBI build 37). The paired-end fragments were unique and located within 400 bp of each other [29].

MethylCap-sequencing analysis

For statistical analysis, reads of promoter (-2000 bp – to + 500 bp of transcription start site) and exon regions were retrieved. In order to identify differences between normal cervices and CIN2/3 lesions, we dichotomised the read data into methylation positive or negative. Samples were considered negative if a sample showed either 0 or 1 read. Samples were considered methylation positive if a sample showed ≥3 reads. Subsequently, regions were ranked based on highest specificity and highest sensitivity for CIN2/3. The candidate markers should fulfil the following criteria: 1) Low/negative reads in the leukocytes to prevent false positive results. The region was excluded if both leukocyte samples showed >1 read or if 1 leukocyte sample showed >2 reads. 2) Unmethylated (0 or 1 read) in at least 75% (15/20) of the normal cervix group. 3) Methylated (≥3 reads) in at least 28% (5/18) of the CIN2/3 lesion group.

Verification and validation of MethylCap-sequencing data by methylation specific PCR (MSP)

MSP primers were designed for the highest ranking top 15 genes (16 DMRs). Sodium bisulfite treatment of isolated genomic DNA (1 µg/sample) was performed according to the recommendations of the EZ DNA methylation kit (Zymo, BaseClear, Leiden, the Netherlands). MSP design and analysis was performed using sequences derived from the H2G2 browser. Each reaction was performed in 30 µl total reaction volume, containing: 600 nM of each MSP primer, 1.5 µl of bisulfite treated DNA (approximately 15 ng), standard PCR components (Applied Biosystems) and 0.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems). Condition of the MSP was: 10 min hot-start at 95°C; 95°C for 60 sec, 60°C for 60 sec, 72°C 60 sec for a total of 40 cycles, with a final elongation step of 7 min at 72°C. Leukocyte DNA from healthy women was used as negative control and *in vitro* methylated (by *SssI* enzyme) leukocyte DNA was used as positive control for each MSP.

Quantitative Methylation Specific PCR (QMSP)

QMSP was performed as described previously by our group with an internal (FAM-ZEN/IBFQ)-labelled hybridisation probe for quantitative analyses [21]. Primer and probe sequences are summarized in Table 1. *β–actin* was used as a methylation

independent internal reference gene. QMSP reactions were performed in 10 µl final volume, containing: 300 nM of forward and reverse primers, 250 nM of hybridisation probe, 5 µl of 2* QuantiTech Probe PCR Master Mix (Qiagen Hilden, Germany) and 2.5 µl bisulfite modified DNA (approximately 25 ng). Each sample was analyzed in triplicate by ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems). Negative and positive controls were the same as used for MSP. Standard curve analysis was performed on each plate and by each primers-probe set on serial dilutions of *in vitro* methylated leukocyte DNA. A DNA sample was considered methylated if at least 2 out of the 3 wells were methylation positive with a Ct-value below 50 and DNA input of at least 225 pg β-actin. The relative level of methylation of the region of interest was determined by the following calculation: the average quantity of the methylated region of interest divided by the average quantity of the reference *β–Actin* gene and multiplied by 10000 [30]. In our analysis we also included 4 genes previously described by our group (*C13ORF18, JAM3, EPB41L3* and *TERT*) to compare sensitivity and specificity of these known genes with the newly identified methylation markers. QMSP for these markers was performed as previously described [21].

HPV testing

HrHPV testing was performed using general primer-mediated PCR (GP5+/6+) as reported previously [30]. For HPV-typing as well as detection of the clinical relevant HPV infections, GP5+/6+ positive cases were tested by COBAS® 4800 HPV test. The COBAS HPV test individually detects HPV 16 and 18, while at the same time identifying 12 additional hrHPV types [31]. The COBAS HPV test is routinely used in our iso-15189-certified laboratory of molecular pathology on scrapings from the national population-based screening program. For the COBAS® HPV testing in this study, the PCR only workflow was used, since no liquid-based scrapings in Preservcyt® were available but only already isolated DNA. This workflow was first validated with DNA isolated from clinical samples that were tested previously in the diagnostic routine and this showed comparable results to the liquid-based samples.

Table 1: MSP Primer and probe sequences of the genes used validation of the top15 genes (16 regions)

Rank	Gene	Region ID		Forward primer 5'->3'	Reverse primer 5'->3'	6FAM/ZEN/IBFQ probe 5'->3'
1	ZSCAN1	7534754	M	TTGTTGGTATTCGTTTGTTC	ACGCGACCGAACGATATT	AGGTCGAAGTTTTTTTACGTATTTTTATTGTTCGTTTA
			U	TTTTTGTTGGTATTTGTTTGTTT	AAACACAACCAAACAATATTT	
2	PCDH17	7180524	M	AGTGAATACGATTCGTTTCG	TAAAAACGCCTCGACATACG	
			U	GAGAGTGAATATGATTTGTTTTGTTG	TCTAAAAACACCTCAACATACAAAA	
3	ST6GALNAC5	7595284	M	GTAGTTGCGGATGGAGGTTC	CTAACTACGCTCACCCTCCG	TTGAAGTTTCGGGTTTGGTCGTCGAGTC
			U	ATTGTAGTTGTGGATGGAGGTTT	ACTAACTACACTCACCCTCCAAT	
4	CLIC6	7721996	M	CGCGTTTTTGGTTGAGAGATTCG	TCGACGCCTCCTCTACG	
			U	TTGTGTTTTTGGTTGAGAGATTTG	CCTCAACACCTCCTCTACAC	
5	AC012354.1	7802984	M	AAATGTATGGGATTTTTGTCGTCG	AAACCCGAACTCCCAACCG	
			U	GTTAAAAATGTATGGGATTTTTGTTGTTG	AAACCCAAACTCCCAACCAC	
6	AL590705.4	8386972	M	TGAGGTTTTATTTGGTTTCGC	TCCGAAACTTAAATACCCGC	
			U	AGTTGAGGTTTTATTTGGTTTTGT	TCCAAAACTTAAATACCCACAAC	
7	PAX2	6998395	M	GATATATAATTAGCGCGGATC	AACGAAACCTCCTAACGCC	CGCGGTTTAGTTTCGGGTATTCGTTTCGG
			U	TTTGTGATATATAATTAGTGTGGATT	AACAAAACCTCCTAACACCCAA	
8	CDH6	8060049	M	GGGCGGCGTTGTTGTC	CCAACCCCACGACGAATC	CGTTTTTCGGGGAGTTTGGGTATCGTTTTTTCG
			U	TGGGTGGTGTTGTTGTT	CCAACCCCACAACAAATC	
9	GFRA1	7006375	M	TAGGGGGAATCGATGTTTC	GAATCCTAAACACCGAACGA	TTTATTCGTCGCGCGTTTTCGG
			U	AGTTAGGGGGAATTGATGTTTT	CCAAATCCTAAACACCAAACAAA	
10	IRX1	8050044	M	TTGTTTTTAGGTTCGGAAAATC	ACCCGTATCCTAAACACCG	
			U	GTTTGTTTTTAGGTTTGGAAAATT	CCACCCATATCCTAAACACCA	
11	POU4F3	8099299	M	GATTTTGGGAAATTCGGTACG	CGCGCGTACAACTATACC	
			U	TTAGATTTTGGGAAATTTGGTAT	CACACATACAACTATACCCCC	
12	GATA4	8293851	M	GGTCGGGTTAATTCGGTC	CCTCGACAAAACTCAAAACG	ATTTCGGTGAGTAGGAGCGCGAG
			U	GGTTGGGTTAATTTGGTT	CCTCAACAAAACTCAAAACA	
13	MKX	6962285	M	AAATTACGCGATAGTAAATCG	AAACCGAAAACATCAAAACG	
14	PAX2	6998393	M	TTTGGAGGGTGCGTAGTC	AATCCAACGAAATCCGAAC	
			U	TTTTTTGGAGGGTGTGTAGTT	AAAATCCAACAAAATCCAAACT	
15	KCNIP4	7992660	M	GGGACGTAGGGTGTAGAAGC	AAACTCTCGCTCCCAACG	TCGGTTAGGGGCGTTTGTTTACGGGTTTGTACGG
			U	TTAGGGATGTAGGGTGTAGAAGT	AAAACTCTCACTCCCAACACC	
16	LHX8	7594866	M	TATTTTTTTCGTAGCGGATC	ACGAAAAACCAAATTCTACG	
			U	TTTTATTTTTTTTGTAGTGGATT	ACAAAAAACCAAATTCTACAACC	

Statistical analysis

Statistical analysis was performed using SPSS software package (SPSS 20, Chicago, IL, USA). Spearman's rank correlation coefficient was used to compare the MethylCap-Seq reads with the MSP band intensity. Categorical methylation data were analyzed using the Pearson χ^2 test. Receiver operating characteristic (ROC) curves were generated and the area under the ROC curve (AUC) was used as a measure of test performance. The Mann-Whitney U test and Kruskall-Wallis test was used to determine differences in methylation ratio in 2 groups or more, respectively. The student T test was used to compare positive methylation and age. To compare sensitivity and specificity of the patient group referred with abnormal cytology by DNA methylation markers versus hrHPV, the extended McNemar test, described by Hawass was executed [32]. P-values lower than 0.05 were considered statistically significant.

Results

Identification of differential methylated genes by MethylCap sequencing

Genome-wide MethylCap-Seq was used to compare the DNA methylation profiles of CIN2/3 dysplastic cervical cells with normal cervical cells to identify CIN2/3 specific DMRs. After applying our criteria, 176 DMRs comprising 163 genes remained. The list of DMRs is shown in supplement table 1, ranked on the sum of unmethylated normal samples and methylated CIN2/3 samples.

Verification and validation of the top15 differentially methylated genes

To verify the MethylCap-Seq data, the top 15, out of the 163 identified genes were selected. MSP primers were designed and could be optimized for 14 out of the 15 genes. Verification of the selected 14 genes showed for 11 genes a significant correlation between the MSP band intensity and the amount of reads from the MethylCap-Seq data. One gene (*PCDH17*) showed high methylation levels in leukocytes and was therefore excluded for further validation. The remaining 10 genes passed verification and continued to the subsequent validation step. Table 2 shows an overview of which genes continued through the different stages of validation.

Table 2: Validation of the top15 genes (16 regions)

Rank	Gene	Region ID	Optimized	Verification	Validation	1st diagnostic evaluation	2nd diagnostic evaluation
1	ZSCAN1	7534754	Yes	Yes	Yes	Yes	Yes
2	PCDH17	7180524	Yes	No*			
3	ST6GALNAC5	7595284	Yes	Yes	Yes	Yes	Yes
4	CLIC6	7721996	Yes	No			
5	AC012354.1	7802984	Yes	No			
6	AL590705.4	8386972	Yes	Yes	Yes	Yes	Yes
7**	PAX2	6998395	Yes	Yes	Yes	No	
8	CDH6	8060049	Yes	Yes	Yes	Yes	Yes
9	GFRA1	7006375	Yes	Yes	Yes	Yes	Yes
10	IRX1	8050044	Yes	No			
11	POU4F3	8099299	Yes	Yes	No*		
12	GATA4	8293851	Yes	Yes	Yes	Yes	Yes
13	MKX	6962285	No				
7**	PAX2	6998393	Yes	No			
14	KCNIP4	7992660	Yes	Yes	Yes	Yes	Yes
15	LHX8	7594866	Yes	Yes	Yes	Yes	Yes

* Excluded due to high methylation in leukocytes
** Same gene, different region

The second validation step was performed by MSP on DNA from FFPE tissue of an independent, randomly selected new patient cohort that consisted of 13 cervical cancers, 19 HSIL lesions (8 CIN2, 8 CIN3 and 3 adCIS) and 17 normal cervices. Out of the 10 genes analyzed, 9 showed low methylation levels in the normal samples, significant differential methylation between normal versus HSIL lesions and again little to no methylation in the leukocytes ($p < 0.05$) (Table 3). These 9 genes (*ZSCAN1, ST6GALNAC5, AL590705.4, PAX2, CDH6, GFRA1, GATA4, KCNIP4* and *LHX8*) were selected for further diagnostic evaluation in cervical scrapings (Table 3).

Table 3: Methylation positivity in an external cohort of FFPE samples to validate results of high methylation in CIN2+ lesions and no methylation in normal cervices of the newly found methylation markers.

Rank	Gene	Normal	CIN 2	CIN3	adCIS	carcinoma
1	ZSCAN1	4/16	8/8	7/8	3/3	12/13
3	ST6GALNAC5	0/16	1/6	4/8	2/3	9/12
6	AL590705.4	0/16	1/8	1/7	2/3	6/12
7	PAX2	1/14	6/8	7/8	3/3	5/13
8	CDH6	1/15	3/8	4/8	3/3	7/13
9	GFRA1	0/12	2/8	3/8	2/3	10/12
11*	POU4F3*	2/14	6/7	3/7	3/3	11/12
12	GATA4	0/17	3/8	2/7	3/3	10/13
14	KCNIP4	0/17	6/8	5/8	3/3	10/12
15	LHX8	1/16	3/8	4/8	3/3	7/13

* Excluded due to high methylation in leukocytes

Diagnostic evaluation by QMSP for normal versus cancer scrapings

To evaluate the diagnostic value of the new methylation markers, cervical scrapings from two cohorts of patients were used: 1) normal versus carcinoma scrapings and 2) scrapings from patients referred from population-based screening with an abnormal Pap smear (≥Pap2). In cohort 1, scrapings of 100 randomly selected cervical carcinoma patients and 89 patients with histologically confirmed normal cervices were used. QMSP analysis showed that the relative levels of DNA methylation were higher in the carcinoma scrapings compared to the normal scrapings for 8 out of the 9 selected genes ($p<0.001$) (Figure 2). The area under the curve (AUC) for methylation ratio in cervical carcinoma showed for 8 genes an AUC>0.91, and for one gene (*PAX2)* an AUC of 0.59. Therefore *PAX2* was excluded from further analysis (Figure 3). In women with a normal cervix, methylation positivity for all 9 genes was not related to age (data not shown).

Diagnostic evaluation by QMSP for normal/LSIL versus HSIL scrapings

In cohort 2, scrapings of 215 consecutive patients referred from population-based screening with an abnormal Pap smear were used. The 8 genes that showed differential methylation in the normal versus the cancer scrapings were subsequently tested in cohort 2. Methylation levels and frequencies for all 8 genes analyzed (*ZSCAN1, ST6GALNAC5, AL590705.4, CDH6, GFRA1, GATA4, KCNIP4* and *LHX8)*, increased with the severity of the underlying histological lesion ($p<0.001$) (Figure 4 and Table 4).

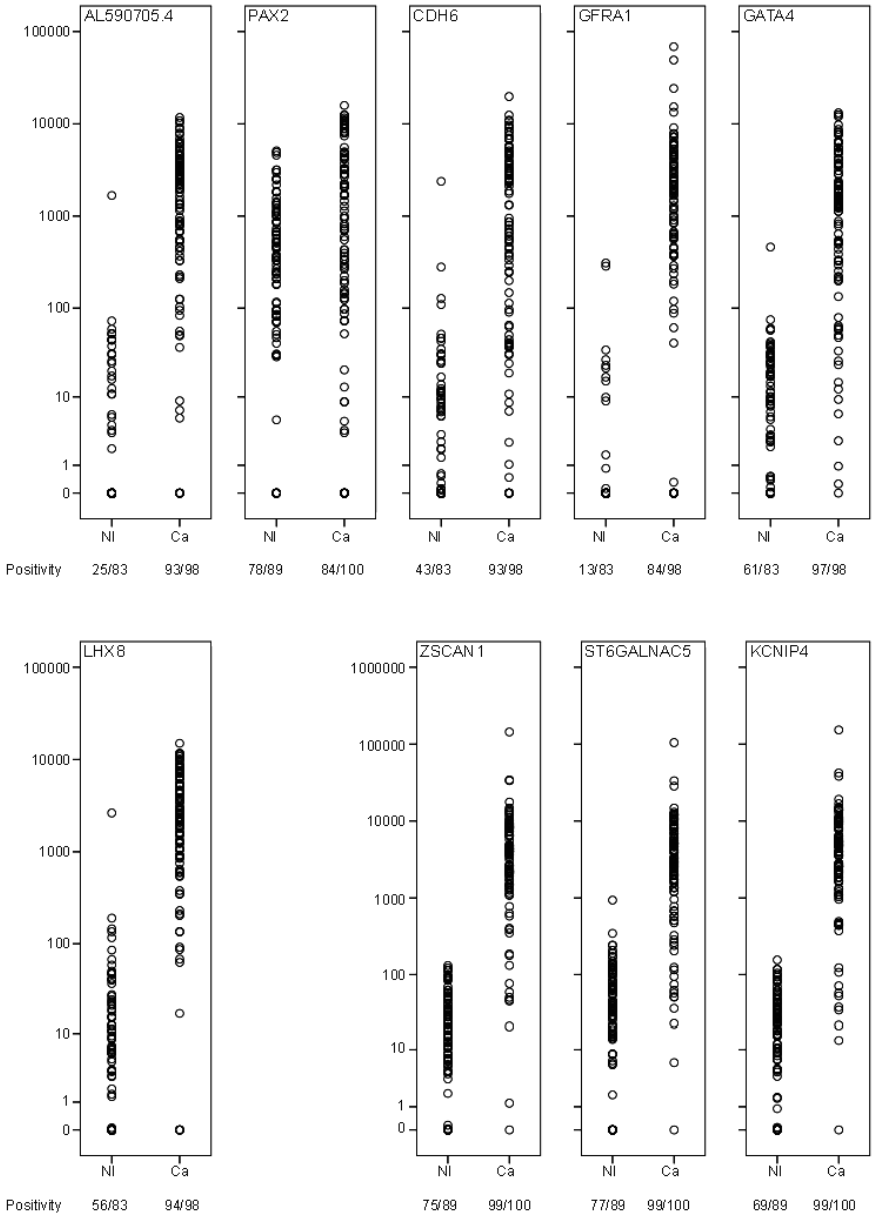

Figure 2: Methylation ratio of 9 genes tested with QMSP in scrapings from normal (Nl) and cancer (Ca) patients. Methylation levels are significantly higher in the cancer scrapings

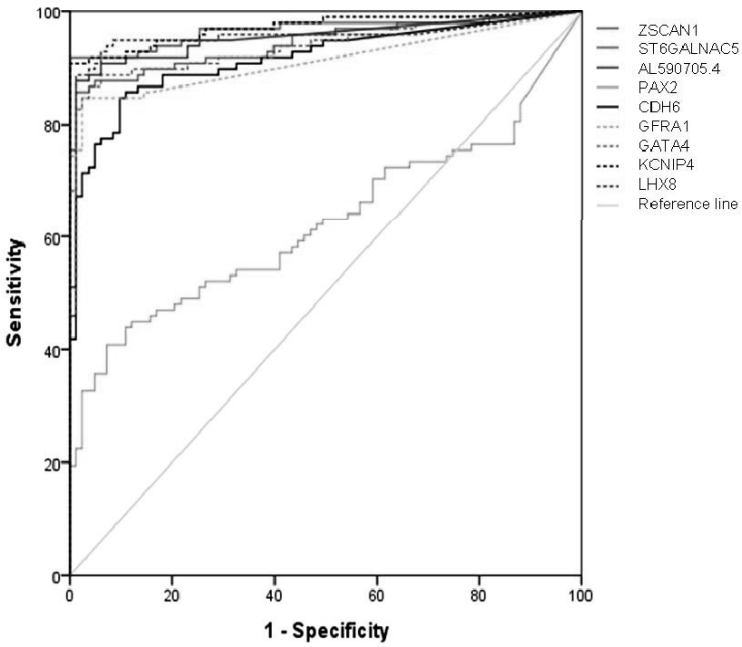

Figure 3: ROC curve analyses of methylation ratios per gene

Without setting a cut-off value for achieving higher/lower sensitivity and/or specificity, genes *ZSCAN1, ST6GALNAC5 and KCNIP4* reached high sensitivity (≥90%) for detection of CIN2+ lesions, while for *CDH6, GATA4* and *LHX8* sensitivity for CIN2+ was between 73-84% (Table 5a). For *AL590705.4* and *GFRA1* sensitivity for CIN2+ was between 46-61%, and these genes showed especially high specificity (82%-92%). In our analysis, we also included a marker panel of 4 genes, previously described by our group (*C13ORF18, JAM3, EPB41L3* and *TERT*) to compare sensitivity and specificity of these known genes with the newly identified methylation markers. The gene *C13ORF18* showed reproducible results as described previously[21] with high specificity (95%) and relatively low sensitivity for CIN2+ of 40%. *JAM3* and *EPB41L3* showed sensitivities for CIN2+ between 63-69% and specificities between 79-91%. The gene *TERT* was previously described with high specificity, but this result could not be reproduced since specificity was only 46% in our analysis, while sensitivity for CIN2+ lesions was 82%.

Table 4. Cytology according to the Papanicolaou system (Bethesda system) per histological subgroup. Methylation and HPV positivity of the 8 new methylation markers and 4 known markers tested with QMSP in cervical scrapings from patients with no CIN, CIN1, CIN2, CIN3 and (mi)Ca (n=215).

Cytology	**No CIN**	**CIN1**	**CIN2**	**CIN3**	**miCA**
Pap2 (ASCUS)	9/27 (33%)	9/38 (24%)	2/49 (4%)	0	0
Pap3A (LSIL)	18/27 (66%)	27/38 (71%)	38/49 (78%)	16/57 (28%)	5/44 (11%)
Pap3B (HSIL)	0	2/38 (5%)	8/49 (16%)	31/57 (54%)	27/44 (61%)
Pap4 (HSIL)	0	0	1/49 (2%)	10/57 (18%)	8/44 (18%)
Pap5 (MiCa)	0	0	0	0	3/44 (7%)
Unknown	0	0	0	0	1/44 (2%)
New genes					
ZSCAN1	20/27 (74%)	28/38 (74%)	45/49 (92%)	51/57 (90%)	44/44 (100%)
ST6GALNAC6	22/27 (82%)	33/38 (87%)	41/49 (84%)	53/57 (93%)	41/44 (93%)
AL590705.4	3/26 (12%)	8/36 (22%)	23/49 (47%)	30/55 (55%)	37/43 (86%)
CDH6	10/26 (39%)	15/36 (42%)	28/49 (57%)	37/55 (67%)	42/43 (98%)
GFRA1	1/26 (4%)	4/36 (11%)	11/49 (22%)	21/55 (38%)	35/43 (81%)
GATA4	14/26 (54%)	21/36 (58%)	38/49 (78%)	44/55 (80%)	41/43 (95%)
KCNIP4	24/27 (89%)	36/38 (95%)	48/49 (98%)	57/57 (100%)	43/44 (98%)
LHX8	12/26 (46%)	20/36 (56%)	33/49 (67%)	44/55 (80%)	41/43 (95%)
Known genes					
C13ORF18	2/27 (7%)	1/38 (3%)	10/49 (20%)	23/57 (40%)	27/44 (61%)
JAM3	3/27 (11%)	3/38 (8%)	21/49 (43%)	36/57 (63%)	37/44 (84%)
EPB41L3	2/27 (7%)	12/38 (32%)	21/49 (43%)	41/57 (72%)	41/44 (93%)
TERT	13/27 (48%)	22/38 (58%)	32/49 (65%)	48/57 (84%)	43/44 (98%)
HPV test					
HrHPV	12/26 (46%)	24/36 (67%)	40/49 (82%)	45/55 (82%)	31/43 (72%)

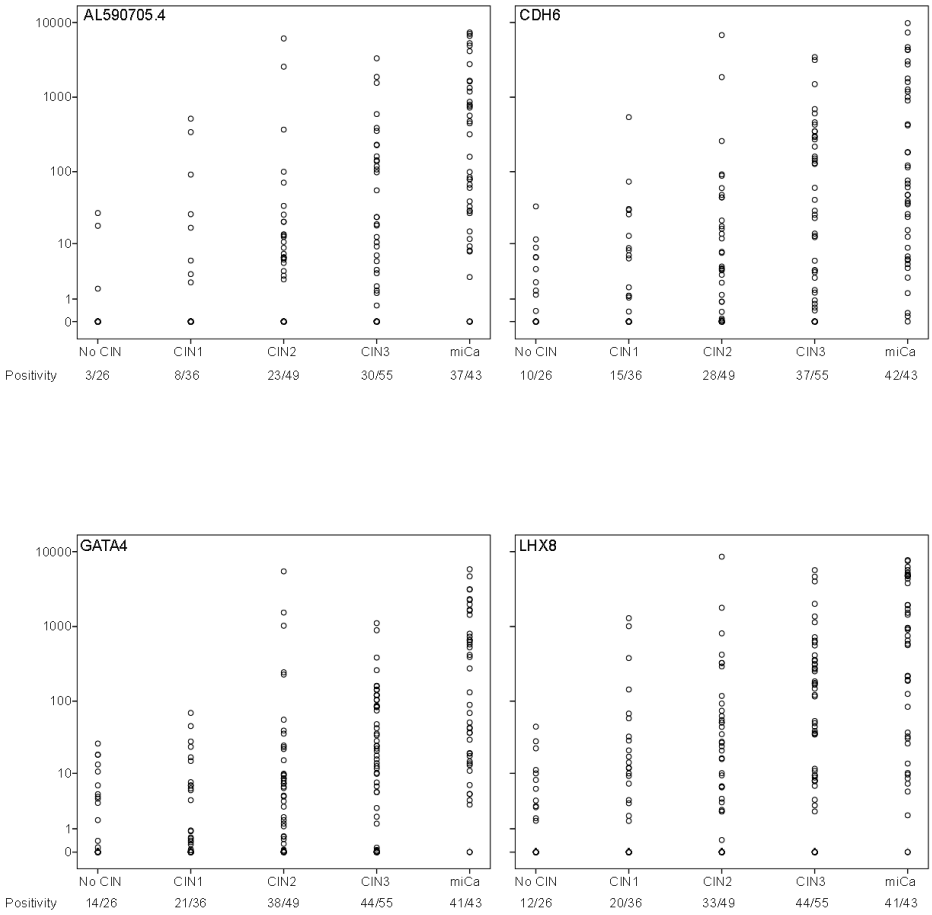

Figure 4: Methylation ratio of the eight genes tested with QMSP in scrapings from patients with No CIN lesion, CIN1, CIN2, CIN3 and (mi)Ca. Relative levels of methylation significantly increases with more severe histological abnormality

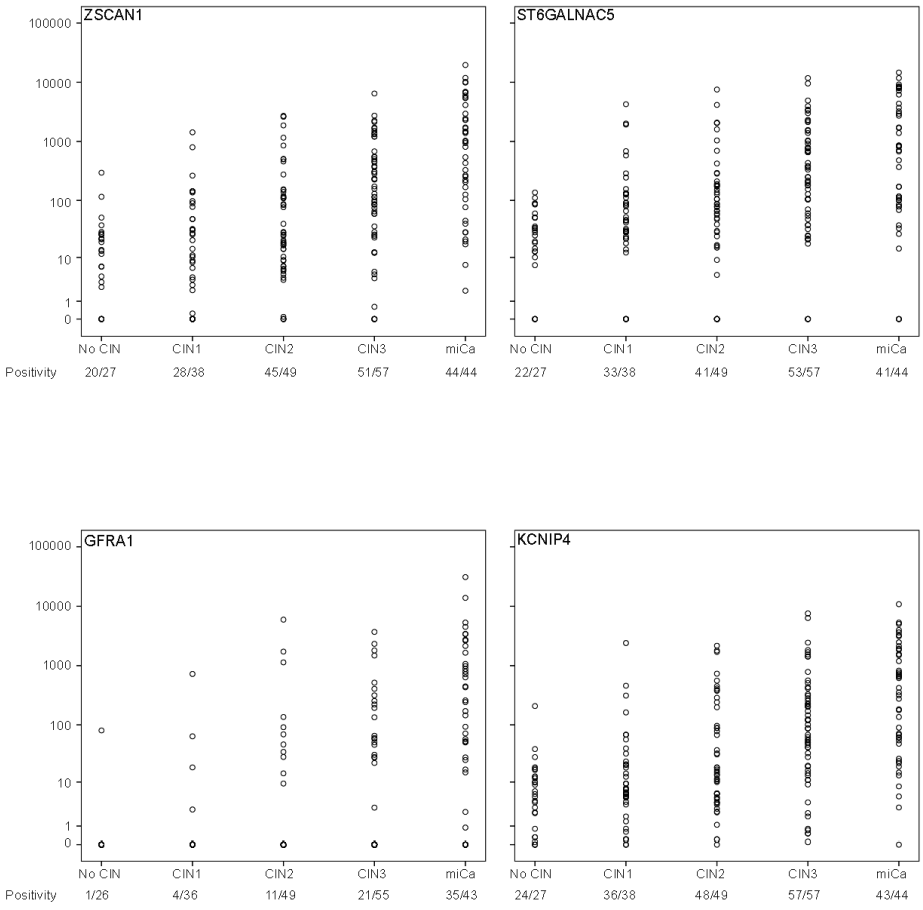

Figure 4: Methylation ratio of the eight genes tested with QMSP in scrapings from patients with No CIN lesion, CIN1, CIN2, CIN3 and (mi)Ca. Relative levels of methylation significantly increases with more severe histological abnormality

hrHPV status and triage testing

HrHPV testing was performed on the patients group referred with abnormal cytology at population-based screening. For 6 out of 215 patients insufficient material was available to perform HPV testing. HrHPV was detected in 152/209 (73%) samples by the GP5+/6+ PCR and COBAS HPV test. Table 4 shows HPV status in relation to underlying histological diagnosis. HrHPV was present in 12/26 (46%) patients without CIN lesion, 24/36 (67%) CIN1 patients, 40/49 (82%) CIN2 patients, 45/55 (82%)

CIN3 patients and 31/43 (72%) patients with miCa. The sensitivity of hrHPV testing for CIN2+ was 79% with a specificity of 42%. For the genes *CDH6, GATA4*, and *LHX8* sensitivity and specificity results were comparable to hrHPV testing with sensitivity for CIN2+ between 73-84% and specificity between 40-60% (Table 5a).

Table 5b shows sensitivity and specificity for CIN2+ and CIN3+ in scrapings of hrHPV positive women (n=152), which were comparable to the results for the whole group, as shown in Table 5a. The genes *ZSCAN1, ST6GALNAC5* and *KCNIP4* again showed high sensitivity (≥92%) for the detection of CIN2+, while for *CDH6, GATA4, EPB41L3, TERT* and *LHX8* sensitivity for CIN2+ was between 72-85%. For *AL590705.4, JAM3, C13ORF18* and *GFRA1* sensitivity for CIN2+ was between 43-68%, however these genes showed high specificity between 86-94%. In the current Dutch population based screening program, women with pap2/pap3a (ASCUS/LSIL) scrapings are retested after 6 months with triage testing by hrHPV. Therefore, we also show the results of triage testing by hrHPV and methylation markers in this group (Table 5b). Triage testing by hrHPV shows a sensitivity for CIN2+ of 82% with a specificity of 41%; *GATA4, LHX8* and *TERT* show comparable results.

Table 5a. Sensitivity and specificity results for CIN2+ and CIN3+ in cervical scrapings from patients referred from population-based screening with an abnormal pap smear (n=215)

Gen	Sensitivity CIN2+	Specificity CIN2+	Sensitivity CIN3+	Specificity CIN3+
ZSCAN1	93%	26%	94%	18%
ST6GALNAC5	90%	15%	93%	16%
AL590705.4	61%	82%	68%	69%
CDH6	73%	60%	81%	52%
GFRA1	46%	92%	57%	86%
GATA4	84%	44%	87%	34%
KCNIP4	99%	8%	99%	5%
LHX8	80%	40%	87%	41%
C13ORF18	40%	95%	50%	87%
EPB41L3	69%	79%	81%	69%
JAM3	63%	91%	72%	76%
TERT	82%	46%	90%	41%
hrHPV	79%	42%	78%	32%

Table 5b. Sensitivity and specificity results for CIN2+ and CIN3+ in scrapings of hrHPV positive women (n=152). And in scrapings of Pap2/Pap3a (ASCUS/LSIL) patients (n=124).

Only hrHPV positive patients (n=152)				
	Sensitivity CIN2+	Specificity CIN2+	Sensitivity CIN3+	Specificity CIN3+
ZSCAN1	94%	36%	96%	22%
ST6GALNAC5	92%	19%	95%	16%
AL590705.4	65%	86%	74%	68%
CDH6	72%	64%	83%	55%
GFRA1	51%	92%	65%	83%
GATA4	85%	47%	88%	33%
KCNIP4	98%	11%	99%	7%
LHX8	81%	53%	91%	45%
C13ORF18	43%	94%	55%	87%
EPB41L3	72%	78%	86%	66%
JAM3	68%	94%	80%	74%
TERT	81%	47%	91%	42%
Only Pap2/3A patients (n=124)				
ZSCAN1	90%	27%	86%	19%
ST6GALNAC5	90%	16%	100%	16%
AL590705.4	38%	82%	40%	74%
CDH6	58%	59%	75%	55%
GFRA1	22%	92%	30%	88%
GATA4	78%	44%	80%	36%
KCNIP4	98%	8%	100%	6%
LHX8	70%	49%	85%	45%
C13ORF18	23%	95%	29%	89%
EPB41L3	51%	78%	76%	72%
JAM3	44%	91%	57%	80%
TERT	74%	48%	91%	43%
hrHPV	82%	41%	80%	32%

Different combinations of genes were analyzed to find the best methylation marker panel with the highest combined sensitivity and specificity. For this analysis a sample was considered positive if either of the genes in the combination tested was positive. By adding more than 3 genes in a combination specificity of the methylation test decreased, with minimal increase in sensitivity. The combinations of genes with the highest combined sensitivity and specificity for CIN2+ was *AL590705.4/EPB41L3/JAM3*

and *AL590705.4/C13ORF18/JAM3* with a sensitivity of 76% and 74%, which is comparable to hrHPV testing (79%). Specificity of both combinations was 71% and 76%, which is significantly higher than for hrHPV testing (42%) (P≤0.05). Table 6 shows that for all other combinations sensitivities for detecting CIN2+ lesions are between 64-80%, with a combined specificity between 58-88%.

Table 6. Combinations of different methylation markers to create a panel of genes most suited as triage test in scrapings, ranked on highest sensitivity (n=215).

Gene combination	Sensitivity CIN2+	Specificity CIN2+	Sensitivity CIN3+	Specificity CIN3+
JAM3/CDH6	80%	58%	86%	47%
AL590705.4/CDH6/EPB41L3	80%	55%	88%	47%
GFRA1/EPB41L3/CDH6	78%	57%	86%	49%
CDH6/EPB41L3	78%	57%	86%	49%
GFRA1/AL590705.4/CDH6	77%	57%	84%	48%
AL590705.4/CDH6	77%	57%	84%	48%
JAM3/EPB41L3/AL590705.4	76%	71%	85%	59%
C13ORF18/JAM3/AL590705.4	74%	76%	81%	60%
GFRA1/EPB41L3/AL590705.4	74%	74%	85%	62%
AL590705.4/EPB41L3	74%	74%	84%	62%
C13ORF18/CDH6	74%	58%	81%	51%
JAM3/GFRA1/AL590705.4	73%	77%	81%	62%
C13ORF18/JAM3/EPB41L3	73%	72%	83%	62%
GFRA1/CDH6	73%	60%	81%	52%
JAM3/AL590705.4	72%	79%	80%	63%
JAM3/EPB41L3	72%	75%	83%	65%
JAM3/EPB41L3/GFRA1	72%	76%	84%	65%
GFRA1/EPB41L3	69%	79%	83%	69%
C13ORF18/EPB41L3	69%	75%	81%	67%
C13ORF18/JAM3/GFRA1	66%	82%	77%	70%
JAM3/GFRA1	65%	86%	76%	73%
C13ORF18/AL590705.4	65%	79%	71%	66%
C13ORF18/JAM3	64%	88%	73%	74%
GFRA1/AL590705.4	64%	81%	71%	68%

Discussion

Due to introduction of primary prevention of cervical cancer through prophylactic vaccination against hrHPV types 16 and 18, involved in 70% of cervical cancer, the incidence of cervical neoplasia will decrease[14]. Current implementation of HPV vaccination programs in Europe will not have a real impact on the incidence of CIN2/3+ within the next 10-15 years. However, this decline in incidence will most probably impair the diagnostic performance of HPV testing and cytology triage testing even more, resulting in less efficient population-based screening programs [14]. There is therefore an urgent need to further improve current methodology for cervical cancer screening.

In this study we report new CIN2/3 specific methylation markers identified by a genome-wide DNA methylation screening strategy comparing CIN2/3 and normal cervical cells. Diagnostic evaluation in cervical scrapings shows that for 8 newly identified genes the relative level of methylation increases with the severity of the underlying histological lesion. Combining our newly identified genes with our previously reported panel (*C13ORF18, JAM3, EPB41L3* and *TERT*) reveals that for the combinations *AL590705.4/EPB41L3/JAM3* and *AL590705.4/C13ORF18/JAM3* sensitivities for CIN2+ are between 74-76%, which is comparable to the sensitivity for CIN2+ of hrHPV testing (79%). Specificities of our gene panel was between 71-76%, which is significantly higher ($p \leq 0.05$) than the specificity for hrHPV testing (42%) in a triage setting after a positive Pap smear test result in population-based screening.

Our strategy identified 163 genes, of which the highest ranking top 15 genes were validated in different steps. From the 163 identified genes, 12 were described previously in literature (*POU4F3, PAX2, WT1, TBX3, SOX1, COL6A2, ALK, SOX17, PCDH10, CTNND2, APOBEC2, hsa-mir-124-1*) as being more frequently methylated in CIN2/3 lesions and/or cervical cancer compared to normal cervices, indicating the validity of our approach. More CIN2+ specific markers are necessary since literature shows that methylation markers were often tested on CIN3 scrapings only [19,20,33].

Cervical cancer screening in the Netherlands in 2016 will change to primary hrHPV screening and because the relatively low specificity of the hrHPV test, a triage test is

necessary for hrHPV positive women to prevent unnecessary referral to the gynecologist. Although triage testing with cytology is now mostly advocated, this test has some disadvantages because of its subjectivity and unreliability to test on self-sampler material [34]. Therefore, we analyzed the performance of our methylation markers in the hrHPV positive scrapings. The combination *AL590705.4/JAM3* and *AL590705.4/C13ORF18/JAM3* showed in the hrHPV positive scrapings sensitivities for CIN2+ between 76-77% and specificities between 81-83% (data not shown). These results are better than for other reported triage strategies in literature, such as immunohistochemical staining with p16INK4a that reports a sensitivity for detecting CIN2/3 around 77% with a specificity of 61% [35,36] or for HPV 16/18 genotyping which reports sensitivity for CIN2/3 around 65% with a specificity of 73% [37]. Since our methylation panel was tested in a selected patient group that was referred from population-based screening with abnormal cytology, further validation in hrHPV positive scrapings collected from a large cohort of women from population-based screening should be performed. These kinds of scrapings from real life cohorts will become available in the Netherlands after 2016 when primary screening has changed to hrHPV testing.

The advantage of methylation analysis is that is an objective test and that it can be performed on the same material used for hrHPV testing, which makes it also interesting for self-sampled material. Different methylation markers already have been tested as a triage test in hrHPV positive women [19,21,33,38-41]. However, for most markers a cut-off value was set in order to obtain high specificity. The advantage of the newly found and previously described markers by our group is that no cut-off value is needed. If the PCR product was negative (i.e. no amplification of specific product), the samples were called negative and any ratio above zero was called positive. This unique feature of the selected genes allows an objective and easy to interpret test.

It is also interesting to speculate about the role of the newly found genes in relation to carcinogenesis. The gene *AL590705.4* is located on chromosome 9 and was not described before as being methylated during carcinogenesis. The gene *CDH6* belongs to the family of Cadherins. Cadherins are membrane glycoproteins that mediate homophilic cell-cell adhesion and play critical roles in cell differentiation and

morphogenesis. Decreased expression of this gene may be associated with tumor growth and metastasis. It has recently been described as a new transforming growth factor-β (TGF-β) target in thyroid tumor patients [42]. The gene *GFRA1* plays a key role in the control of neuron survival and differentiation. It has been described as differentally methylated between cancerous and non-cancerous tissue obtained from lung cancer patients based on DNA methylation profiles [43].

The strengths of our current study are: 1) the genome-wide approach with MethylCap-Seq for specific identification of differential methylation regions between normal cervices and CIN2/3 lesions, 2) the systematic verification and validation of the markers found, and 3) the selection of the best performing markers for diagnostic evaluation in cervical scrapings. The limitation of the current study was that diagnostic evaluation of the markers was performed on a selected group of patients that were referred to our hospital for abnormal cytological screening and not on HPV positivity since that is not allowed by law in the Netherlands (unless within clinical trials). The sensitivity of the cytological screening test is lower than for hrHPV screening, therefore some CIN2+ cases that are hrHPV positive, will be cytological normal. We cannot exclude that inclusion of such cases might affect the results. This is a target for future studies.

Population-based screening is in transition and methylation markers might be an important component in future screening settings. It is important to validate the most interesting markers described in literature in a population-based screening trial. Verification of the results by different groups is important to assure the reproducibility of the methylation analysis. The combination of genes with the highest possible sensitivity and specificity should be evaluated.

In conclusion, we identified new CIN2/3 specific methylation markers for detection of cervical neoplasia in cervical scrapings. These newly found markers might be applied as a triage test in hrHPV positive women from population-based screening. However, further validation in population-based cohorts is needed.

References

1. Peto J, Gilham C, Fletcher O, Matthews FE. The cervical cancer epidemic that screening has prevented in the UK. *Lancet*. 2004;364(9430):249-256.

2. Arbyn M, Raifu AO, Weiderpass E, Bray F, Anttila A. Trends of cervical cancer mortality in the member states of the european union. *Eur J Cancer*. 2009;45(15):2640-2648.

3. Cuzick J, Clavel C, Petry KU, et al. Overview of the european and north american studies on HPV testing in primary cervical cancer screening. *Int J Cancer*. 2006;119(5):1095-1101.

4. Mayrand MH, Duarte-Franco E, Rodrigues I, et al. Human papillomavirus DNA versus papanicolaou screening tests for cervical cancer. *N Engl J Med*. 2007;357(16):1579-1588.

5. Cox JT, Castle PE, Behrens CM, et al. Comparison of cervical cancer screening strategies incorporating different combinations of cytology, HPV testing, and genotyping for HPV 16/18: Results from the ATHENA HPV study. *Am J Obstet Gynecol*. 2013;208(3):184.e1-184.e11.

6. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol*. 1999;189(1):12-19.

7. Ronco G, Dillner J, Elfstrom KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: Follow-up of four european randomised controlled trials. *Lancet*. 2013.

8. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: A randomised controlled trial. *Lancet Oncol*. 2010;11(3):249-257.

9. Bulkmans NW, Berkhof J, Rozendaal L, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *Lancet*. 2007;370(9601):1764-1772.

10. Rijkaart DC, Berkhof J, Rozendaal L, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: Final results of the POBASCAM randomised controlled trial. *Lancet Oncol*. 2012;13(1):78-88.

11. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine*. 2012;30 Suppl 5:F88-99.

12. Cuzick J, Arbyn M, Sankaranarayanan R, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. *Vaccine*. 2008;26 Suppl 10:K29-41.

13. Kulasingam SL, Hughes JP, Kiviat NB, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: Comparison of sensitivity, specificity, and frequency of referral. *JAMA*. 2002;288(14):1749-1757.

14. Franco EL, Mahmud SM, Tota J, Ferenczy A, Coutlee F. The expected impact of HPV vaccination on the accuracy of cervical cancer screening: The need for a paradigm change. *Arch Med Res*. 2009;40(6):478-485.

15. Baylin SB, Ohm JE. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer*. 2006;6(2):107-116.

16. Steenbergen RD, Snijders PJ, Heideman DA, Meijer CJ. Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nat Rev Cancer*. 2014;14(6):395-405.

17. Yang N, Nijhuis ER, Volders HH, et al. Gene promoter methylation patterns throughout the process of cervical carcinogenesis. *Cell Oncol*. 2010;32(1-2):131-143.

18. Bierkens M, Hesselink AT, Meijer CJ, et al. CADM1 and MAL promoter methylation levels in hrHPV-positive cervical scrapes increase proportional to degree and duration of underlying cervical disease. *Int J Cancer*. 2013;133(6):1293-1299.

19. Hesselink AT, Heideman DA, Steenbergen RD, et al. Combined promoter methylation analysis of CADM1 and MAL: An objective triage tool for high-risk human papillomavirus DNA-positive women. *Clin Cancer Res*. 2011;17(8):2459-2465.

20. Lai HC, Lin YW, Huang RL, et al. Quantitative DNA methylation analysis detects cervical intraepithelial neoplasms type 3 and worse. *Cancer*. 2010;116(18):4266-4274.

21. Eijsink JJ, Lendvai A, Deregowski V, et al. A four-gene methylation marker panel as triage test in high-risk human papillomavirus positive patients. *Int J Cancer*. 2012;130(8):1861-1869.

22. Yang N, Eijsink JJ, Lendvai A, et al. Methylation markers for CCNA1 and C13ORF18 are strongly associated with high-grade cervical intraepithelial neoplasia and cervical cancer in cervical scrapings. *Cancer Epidemiol Biomarkers Prev*. 2009;18(11):3000-3007.

23. Reesink-Peters N, Wisman GB, Jeronimo C, et al. Detecting cervical cancer by quantitative promoter hypermethylation assay on cervical scrapings: A feasibility study. *Mol Cancer Res*. 2004;2(5):289-295.

24. Laird PW. Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet*. 2010;11(3):191-203.

25. Lendvai A, Johannes F, Grimm C, et al. Genome-wide methylation profiling identifies hypermethylated biomarkers in high-grade cervical intraepithelial neoplasia. *Epigenetics*. 2012;7(11):1268-1278.

26. Rauch TA, Pfeifer GP. DNA methylation profiling using the methylated-CpG island recovery assay (MIRA). *Methods*. 2010;52(3):213-217.

27. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 concerted action BMH4-CT98-3936. *Leukemia*. 2003;17(12):2257-2317.

28. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25-2009-10-3-r25. Epub 2009 Mar 4.

29. De Meyer T, Mampaey E, Vlemmix M, et al. Quality evaluation of methyl binding domain based kits for enrichment DNA-methylation sequencing. *PLoS One*. 2013;8(3):e59068.

30. Wisman GB, Nijhuis ER, Hoque MO, et al. Assessment of gene promoter hypermethylation for detection of cervical neoplasia. *Int J Cancer*. 2006;119(8):1908-1914.

31. Cui M, Chan N, Liu M, et al. Clinical performance of roche cobas 4800 HPV test. *J Clin Microbiol*. 2014;52(6):2210-2211.

32. Hawass NE. Comparing the sensitivities and specificities of two diagnostic procedures performed on the same group of patients. *Br J Radiol*. 1997;70(832):360-366.

33. Overmeer RM, Louwers JA, Meijer CJ, et al. Combined CADM1 and MAL promoter methylation analysis to detect (pre-)malignant cervical lesions in high-risk HPV-positive women. *Int J Cancer*. 2011;7(6).

34. Gok M, van Kemenade FJ, Heideman DA, et al. Experience with high-risk human papillomavirus testing on vaginal brush-based self-samples of non-attendees of the cervical screening program. *Int J Cancer*. 2012;130(5):1128-1135.

35. Carozzi F, Confortini M, Dalla Palma P, et al. Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: A nested substudy of the NTCC randomised controlled trial. *Lancet Oncol*. 2008;9(10):937-945.

36. Carozzi F, Gillio-Tos A, Confortini M, et al. Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: A prospective analysis of a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol*. 2013;14(2):168-176.

37. Rijkaart DC, Berkhof J, van Kemenade FJ, et al. Evaluation of 14 triage strategies for HPV DNA-positive women in population-based cervical screening. *Int J Cancer*. 2012;130(3):602-610.

38. Brentnall AR, Vasiljevic N, Scibior-Bentkowska D, et al. A DNA methylation classifier of cervical precancer based on human papillomavirus and human genes. *Int J Cancer*. 2014.

39. Vasiljevic N, Scibior-Bentkowska D, Brentnall AR, Cuzick J, Lorincz AT. Credentialing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women. *Gynecol Oncol*. 2014;132(3):709-714.

40. Hansel A, Steinbach D, Greinke C, et al. A promising DNA methylation signature for the triage of high-risk human papillomavirus DNA-positive women. *PLoS One*. 2014;9(3):e91905.

41. Verhoef VM, Bosgraaf RP, van Kemenade FJ, et al. Triage by methylation-marker testing versus cytology in women who test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): A randomised controlled non-inferiority trial. *Lancet Oncol*. 2014;15(3):315-322.

42. Sancisi V, Gandolfi G, Ragazzi M, et al. Cadherin 6 is a new RUNX2 target in TGF-beta signalling pathway. *PLoS One*. 2013;8(9):e75489.

43. Selamat SA, Chung BS, Girard L, et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. *Genome Res*. 2012;22(7):1197-1211.

Supplement table 1: Identified DMRs with sensitivity and specificity ranked on the sum of both

Rank	Region ID	Gene	Size region	Nr normals unmethylated	Nr CIN2/3 methylated	Sum unmeth normals and meth CIN2/3
1	7534754	ZSCAN1	195	19	9	28
2	7180524	PCDH17	133	19	8	27
3	7595284	ST6GALNAC5	180	19	8	27
4	7721996	CLIC6	123	20	6	26
5	7802984	AC012354.1	262	17	9	26
6	8386972	AL590705.4	412	17	9	26
7	6998395	PAX2	142	20	5	25
8	8060049	CDH6	139	20	5	25
9	7006375	GFRA1	313	19	6	25
10	8050044	IRX1	171	19	6	25
11	8099299	POU4F3	270	18	7	25
12	8293851	GATA4	264	18	7	25
13	6962285	MKX	436	17	8	25
14	6998393	PAX2	427	17	8	25
15	7992660	KCNIP4	305	17	8	25
16	7594866	LHX8	443	16	9	25
17	7610478	RP11-439A17.1	245	15	10	25
18	6998398	PAX2	197	19	5	24
19	7019754	DPYSL4	246	19	5	24
20	7037878	WT1	198	19	5	24
21	8000998	GABRA2	151	19	5	24
22	7319687	CACNG3	70	18	6	24
23	8021385	TNIP3	126	18	6	24
24	8323242	ZFHX4	282	18	6	24
25	6998396	PAX2	386	17	7	24
26	7256161	GJD2	527	17	7	24
27	7633581	AL136370.1	1	17	7	24
28	8070077	RGS7BP	339	17	7	24
29	8173181	RLBP1L2	223	17	7	24

30	8371278	C9orf71	41	17	7	24
31	7186432	SLITRK1	402	15	9	24
32	8036518	GALNTL6	325	15	9	24
33	8444752	FAM133A	264	15	9	24
34	7144043	TBX3	120	16	8	24
35	7198746	SOX1	432	16	8	24
36	7650220	EP11-10E13.1.1	138	16	8	24
37	7530232	CACNG7	343	15	8	23
38	7531604	EPS8L1	364	15	8	23
39	7533332	ZNF542	400	15	8	23
40	8389207	GRIN3A	372	15	8	23
41	7035376	SLC6A5	422	16	7	23
42	7262409	SHF	272	16	7	23
43	7614837	TCHH	160	16	7	23
44	7733864	COL6A2	317	16	7	23
45	7939583	AC092691.1	273	16	7	23
46	8006035	EPHA5	260	16	7	23
47	8028522	POU4F2	443	16	7	23
48	8153679	TFAP2B	213	16	7	23
49	8299826	PEBP4	55	16	7	23
50	7037882	WT1	195	18	5	23
51	7319688	CACNG3	206	18	5	23
52	8420025	MXRA5	166	18	5	23
53	7015183	FOXI2	371	17	6	23
54	7198756	SOX1	108	17	6	23
55	7500952	ZNF676	299	17	6	23
56	7534399	ZNF154	168	17	6	23
57	7795180	ALK	451	17	6	23
58	7853759	KCNH7	240	17	6	23
59	7932344	CADM2	159	17	6	23
60	7954401	ZIC4	331	17	6	23
61	8220882	ELMO1	229	17	6	23
62	8300584	NKX2-6	198	17	6	23

63	8308835	ZMAT4	239	17	6	23
64	8354231	FOXH1	42	17	6	23
65	6997382	NKX2-3	468	15	7	22
66	7209496	SSTR1	224	15	7	22
67	7300000	BAIAP3	425	15	7	22
68	7488800	ZNF69	181	15	7	22
69	7949963	PPP2R3A	1	15	7	22
70	7950795	AC007159.2	554	15	7	22
71	8031815	NPY2R	512	15	7	22
72	8095341	5S_rRNA	192	15	7	22
73	8144463	AL033519.2	1	15	7	22
74	8161319	SNAP91	466	15	7	22
75	8161598	TBX18	569	15	7	22
76	8315432	SOX17	137	15	7	22
77	8369935	BX248098.1	160	15	7	22
78	8444558	PCDH11X	1	15	7	22
79	6973216	DRGX	516	16	6	22
80	7168291	PDX1	259	16	6	22
81	7216677	C14orf39	288	16	6	22
82	7558688	VWA5B1	16	16	6	22
83	7721995	CLIC6	143	16	6	22
84	7739980	PI1KAP2	144	16	6	22
85	8136141	NRSN1	147	16	6	22
86	8166500	C6orf220	206	16	6	22
87	8209902	THSD7A	229	16	6	22
88	8355044	FOXD4	384	16	6	22
89	7051745	KCNK4	388	17	5	22
90	7096188	KCNA6	207	17	5	22
91	7135359	C12orf42	257	17	5	22
92	7172465	C13orf36	548	17	5	22
93	7216685	SIX6	226	17	5	22
94	7255970	NOP10	1	17	5	22
95	7399713	AC091132.3	1	17	5	22

96	7429544	CBX2	219	17	5	22
97	7528453	ZNF578	132	17	5	22
98	7582721	MKNK1	1	17	5	22
99	7584213	DMRTA2	203	17	5	22
100	7769721	FAM19A5	307	17	5	22
101	7810978	OTX1	198	17	5	22
102	7848584	AC064865.1	1	17	5	22
103	7858997	HOXD13	135	17	5	22
104	8017441	DKK2	272	17	5	22
105	8136976	SCGN	312	17	5	22
106	8169072	MICAL1	1	17	5	22
107	8318803	BHLHE22	461	17	5	22
108	7031638	PARVA	138	15	6	21
109	7150264	AC084018.1	111	15	6	21
110	7251344	GABRG3	423	15	6	21
111	7255084	GREM1	426	15	6	21
112	7255394	RYR3	378	15	6	21
113	7489068	ZNF625	1	15	6	21
114	7589242	DAB1	426	15	6	21
115	7602550	AL136147.1	227	15	6	21
116	7613795	FAM63A	36	15	6	21
117	7689874	SLC32A1	191	15	6	21
118	7720644	OLIG1	308	15	6	21
119	7778034	SOX11	121	15	6	21
120	7843418	AC140481.1	354	15	6	21
121	7851097	GALNT13	179	15	6	21
122	7870465	PTH2R	337	15	6	21
123	7946628	KIAA1257	101	15	6	21
124	8024629	PCDH10	227	15	6	21
125	8218592	AVL9	63	15	6	21
126	8334444	RSPO2	402	15	6	21
127	6973214	DRGX	253	16	5	21
128	6998512	SEMA4G	1	16	5	21

129	7035410	NELL1	115	16	5	21
130	7216686	SIX6	218	16	5	21
131	7473720	ONECUT3	506	16	5	21
132	7584206	DMRTA2	208	16	5	21
133	7619450	ETV3L	55	16	5	21
134	7645452	AC011700.2	291	16	5	21
135	7681392	GINS1	80	16	5	21
136	7893422	CNTN4	406	16	5	21
137	7946826	RAB43	7	16	5	21
138	8054394	CTNND2	321	16	5	21
139	8096965	PCDHGC5	233	16	5	21
140	8165363	AL137784.3	160	16	5	21
141	8237177	WBSCR17	323	16	5	21
142	8252950	AC006329.2	494	16	5	21
143	8300582	NKX2-6	217	16	5	21
144	8411390	GBGT1	201	16	5	21
145	6999108	FBXW4	100	15	5	20
146	6999851	PSD	325	15	5	20
147	7028932	TUB	120	15	5	20
148	7096257	KCNA1	338	15	5	20
149	7101957	GRIN2B	405	15	5	20
150	7209031	NKX2-1	137	15	5	20
151	7223639	ZFYVE1	63	15	5	20
152	7327883	MYLK3	99	15	5	20
153	7394459	KRT27	1	15	5	20
154	7398204	RUNDC3A	314	15	5	20
155	7528533	ZNF808	419	15	5	20
156	7552927	RP1-21O18.1	34	15	5	20
157	7614845	TCHH	330	15	5	20
158	7621232	TOMM40L	140	15	5	20
159	7692820	GDAP1L1	374	15	5	20
160	7762360	TTLL12	134	15	5	20
161	7825980	CNGA3	170	15	5	20

162	8000034	PHOX2B	221	15	5	20
163	8018500	ELOVL6	27	15	5	20
164	8022574	FAT4	369	15	5	20
165	8031643	DCHS2	62	15	5	20
166	8096823	PCDHB15	233	15	5	20
167	8147247	DNAH8	224	15	5	20
168	8148415	APOBEC2	50	15	5	20
169	8199655	UNC84A	190	15	5	20
170	8247593	CASD1	1	15	5	20
171	8287434	DLGAP2	391	15	5	20
172	8292161	hsa-mir-124-1	135	15	5	20
173	8293849	GATA4	179	15	5	20
174	8301460	EBF2	186	15	5	20
175	8413181	AL354796.1	379	15	5	20
176	8430345	BX842568.1	7	15	5	20